

BAB I

PENDAHULUAN

1.1. Latar Belakang

Ketahanan pangan dan gizi merupakan pilar fundamental dalam pembangunan sumber daya manusia yang berkualitas di seluruh dunia (Watson dkk., 2019). Indonesia sebagai negara berkembang menghadapi tantangan serius dalam hal ketahanan gizi dan pembangunan sumber daya manusia. Salah satu permasalahan utama yang dihadapi adalah *stunting* atau kondisi gagal tumbuh pada anak akibat kekurangan gizi kronis. Berdasarkan data Studi Status Gizi Indonesia prevalensi *stunting* di Indonesia berada pada angka 21,6%, masih berada di atas ambang batas yang ditetapkan oleh World Health Organization (WHO) sebesar 20% (Wibisono & Santosa, 2026). Kondisi ini tidak hanya mempengaruhi pertumbuhan fisik anak, tetapi juga berdampak pada perkembangan kognitif serta produktivitas mereka di masa depan (Mediani dkk., 2020).

Sebagai upaya mengatasi permasalahan gizi dan *stunting* tersebut, pemerintah Indonesia melalui Badan Gizi Nasional (BGN) meluncurkan program Makan Bergizi Gratis (MBG) yang resmi dilaksanakan pada 6 Januari 2025 sebagai salah satu program prioritas dalam Asta Cita pemerintahan Prabowo Subianto - Gibran Rakabuming Raka (Aisya, 2025). Program MBG menyasar peserta didik mulai dari jenjang PAUD hingga SMA/ sederajat, balita, ibu hamil, dan ibu menyusui. Skema penyediaan makanan dilakukan secara langsung di sekolah dan posyandu (Purwowidhu, 2025). Serta, memanfaatkan bahan pangan lokal untuk mendukung kesejahteraan petani dan pelaku UMKM. Implementasi program ini diharapkan tidak hanya mengatasi masalah gizi dan *stunting*, tetapi juga meningkatkan kehadiran siswa dan konsentrasi belajar (Lendra dkk., 2025).

Hingga Oktober 2025, program MBG telah menjangkau lebih dari 36 juta penerima manfaat yang tersebar di berbagai wilayah Indonesia. Lebih dari 1,6 juta pekerja langsung dan 2,5 juta pekerja tidak langsung terlibat dalam ekosistem MBG. MBG dirancang sebagai salah satu instrumen strategis untuk menekan angka kemiskinan dan mengurangi ketimpangan antarwilayah (Badan Gizi Nasional, 2025).

Program MBG sejalan dengan tujuan dalam *Sustainable Development*

Goals (SDGs), khususnya SDG 2 yaitu *Zero Hunger* dengan meningkatkan akses terhadap makanan bergizi, SDG 3 yaitu *Good Health and Well-being* melalui upaya perbaikan status gizi dan kesehatan masyarakat, serta SDG 4 yaitu *Quality Education* melalui dukungan terhadap konsentrasi belajar dan kualitas pendidikan peserta didik (United Nations Indonesia, 2026).

Meskipun memiliki tujuan yang strategis, implementasi program MBG tidak terlepas dari berbagai tanggapan publik. Program ini memperoleh dukungan karena dinilai mampu meningkatkan akses terhadap makanan bergizi bagi masyarakat rentan, mendukung upaya penurunan *stunting*, serta memberikan dampak ekonomi melalui penciptaan lapangan kerja dan keterlibatan pelaku usaha lokal dalam rantai pasok pangan. Namun, dalam pelaksanaannya program ini juga menghadapi sejumlah kritik yang berkaitan dengan kualitas makanan, transparansi pengelolaan anggaran, serta besarnya alokasi dana yang mencapai Rp71 triliun dalam APBN 2025 (Umi, 2025). Besarnya anggaran tersebut memunculkan perdebatan di ruang publik mengenai efektivitas penggunaan dana negara dan prioritas program pemerintah dibandingkan kebutuhan pembangunan di sektor lainnya. Selain itu, berdasarkan data Kementerian Kesehatan RI, hingga 5 Oktober 2025 tercatat hampir 12 ribu kasus keracunan yang terkait dengan program MBG (Kautsar, 2025). Kondisi ini mendorong munculnya berbagai opini publik yang banyak disampaikan melalui media sosial.

Berbagai opini berupa kritik maupun dukungan terhadap program MBG banyak disampaikan publik melalui media sosial. Sebagai *platform* komunikasi digital, media sosial memungkinkan pengguna menyampaikan opini secara cepat dan *real-time*, sehingga menjadi sumber data yang kaya untuk mengamati opini publik (Wahabi & Febriana, 2022). Salah satu *platform* yang banyak digunakan untuk mendiskusikan berbagai isu publik adalah media sosial X (Twitter). Berdasarkan laporan *Global Overview* oleh *We Are Social Digital* (2024), jumlah pengguna aktif media sosial X di Indonesia mencapai 139 juta jiwa atau setara dengan 49,9% dari total populasi Indonesia. Kondisi tersebut menjadikan X sebagai sumber data yang relevan untuk menganalisis opini publik terhadap program MBG.

Untuk memahami emosi yang terkandung dalam *tweet*, penelitian ini

menggunakan analisis multikelas emosi yang memungkinkan peneliti memahami emosi yang terkandung dalam *tweet* (Khadapi dkk., 2023). Hal ini sangat penting untuk memperoleh gambaran yang jelas mengenai pandangan publik terhadap isu-isu yang berkaitan dengan MBG (Vindua dkk., 2023). Pendekatan tradisional dalam analisis opini publik umumnya terbatas pada klasifikasi emosi biner, yaitu positif dan negatif. Sedangkan, dalam konteks program yang bersifat kompleks dan menyentuh aspek emosional publik, pendekatan klasifikasi emosi biner belum mampu menggambarkan variasi emosi masyarakat. Sebagai contoh, emosi marah dan takut keduanya memiliki polaritas negatif, tetapi memiliki implikasi yang benar-benar berbeda (Sondakh dkk., 2023), sehingga diperlukan pendekatan yang lebih komprehensif melalui analisis emosi multikelas.

Dalam kajian emosi, terdapat beberapa teori yang digunakan untuk merepresentasikan struktur emosi manusia. Russell (1980) mengembangkan *Circumplex Model of Affect* yang merepresentasikan emosi secara dimensional berdasarkan dua dimensi utama, yaitu *valence* yang menunjukkan tingkat kesenangan atau ketidaksenangan dan *arousal* yang menunjukkan tingkat aktivasi emosional. Sementara itu, Plutchik (1980) mengembangkan teori psiko-evolusioner yang mengelompokkan emosi ke dalam delapan emosi dasar, yaitu senang (*joy*), percaya (*trust*), takut (*fear*), terkejut (*surprise*), sedih (*sadness*), jijik (*disgust*), marah (*anger*), dan antisipasi (*anticipation*), yang dapat memiliki tingkat intensitas dan membentuk kombinasi emosi yang lebih kompleks. Berbeda dengan pendekatan tersebut, Ekman (1992) mengemukakan pendekatan emosi dasar yang bersifat diskrit, yaitu pendekatan yang menganggap emosi sebagai kategori yang terpisah, memiliki karakteristik tersendiri, dan bukan rentang kontinu. Dengan membedakan emosi ke dalam enam kategori, yaitu senang, sedih, marah, takut, jijik, dan terkejut.

Penelitian ini menggunakan teori emosi dasar Ekman yang sejalan dengan tujuan penelitian yaitu klasifikasi multikelas dengan menetapkan setiap *tweet* ke dalam label emosi yang berbeda secara diskrit. Berbeda dengan pendekatan dimensional Russell yang merepresentasikan emosi dalam ruang kontinu, pendekatan ini memandang emosi sebagai suatu spektrum atau rentang nilai berdasarkan dimensi tertentu, sehingga suatu emosi dapat berada pada tingkat

intensitas yang berbeda dan tidak terbatas pada kategori yang terpisah. Kategori emosi Ekman memiliki jumlah kelas yang lebih terstruktur dibandingkan model Plutchik. Dengan demikian, penggunaan enam emosi dasar Ekman memungkinkan respon publik terhadap program MBG direpresentasikan lebih spesifik dibandingkan klasifikasi biner positif dan negatif.

Dalam statistika, permasalahan klasifikasi multikelas dapat dianalisis menggunakan regresi logistik multinomial, yaitu metode yang digunakan untuk mengklasifikasikan data ke dalam lebih dari dua kategori berdasarkan karakteristik data tersebut. Namun, regresi logistik multinomial memiliki keterbatasan dalam menangani data teks berdimensi tinggi karena memerlukan proses ekstraksi fitur secara manual dan umumnya menggunakan representasi kata yang belum mempertimbangkan hubungan kontekstual antar kata (Aggarwal, 2012). Selain itu, model berbasis fitur tradisional memiliki keterbatasan dalam memahami makna kata yang bergantung pada konteks kalimat, terutama pada teks tidak terstruktur seperti *tweet* yang mengandung variasi bahasa, singkatan, dan ekspresi informal (Yung dkk., 2018). Kompleksitas tersebut menyebabkan identifikasi emosi tidak efektif apabila dilakukan secara manual dan memerlukan pendekatan komputasional yang mampu mempelajari pola bahasa, serta mengolah data dalam skala besar secara efisien. Oleh karena itu, penelitian ini menggunakan pendekatan *machine learning* untuk melakukan klasifikasi emosi pada data teks.

Machine learning merupakan komponen utama *Artificial Intelligence* untuk menyelesaikan berbagai permasalahan (Roihan dkk., 2019). Penggunaan *Machine Learning* bertujuan untuk mengajarkan mesin agar dapat memproses data lebih efisien (Mahesh, 2018). Melalui pengidentifikasian pola dalam dataset, *Machine Learning* mampu memprediksi dan memahami karakteristik dari objek yang tidak dikenal (Ula dkk., 2021). Dalam *machine learning*, proses klasifikasi dilakukan melalui dua pendekatan utama, yaitu *supervised learning* dan *unsupervised learning*. Dalam *Supervised Learning*, model belajar menggunakan data yang sudah diberi label (Mestika dkk., 2022), sedangkan pada *Unsupervised Learning* model harus menemukan pola dari data tanpa label (Montavon dkk., 2022). Penelitian ini menggunakan pendekatan *supervised learning*, karena setiap data *tweet* terlebih dahulu diberi label emosi, sebelum dilakukan proses pelatihan

model.

Dalam implementasinya, pendekatan *supervised learning* memerlukan metode yang mampu mempelajari hubungan antara data masukan dan keluaran secara efektif. Salah satu metode yang banyak digunakan adalah *Artificial Neural Network* (ANN) atau Jaringan Syaraf Tiruan, yang terinspirasi dari cara kerja sistem saraf biologis manusia dalam memproses informasi. ANN terdiri dari sejumlah *neuron* yang tersusun dalam beberapa lapisan (*layer*) untuk memodelkan hubungan antara data masukan dan keluaran (Haykin, 2009). Meskipun ANN konvensional telah banyak digunakan dalam berbagai permasalahan klasifikasi, model ini masih memiliki keterbatasan dalam menangkap representasi data yang sangat kompleks, terutama pada data dengan dimensi tinggi dan pola yang tidak terstruktur seperti *tweet* pada media sosial X. Oleh karena itu, berkembanglah pendekatan *Deep Learning* yang mengadaptasi pendekatan *multilayer* ke lapisan *neural network* (Dang dkk., 2020). Disebut *deep* karena model ini terdiri dari banyak lapisan tersembunyi (*hidden layer*) yang memungkinkan pembelajaran representasi data secara bertingkat, sehingga mampu menangkap pola yang lebih kompleks dibandingkan metode *machine learning* konvensional (Nurhalizah dkk., 2024). Pendekatan ini terbukti lebih efektif dalam menangani data yang kompleks, termasuk data teks dalam jumlah besar seperti yang dihasilkan oleh media sosial X.

Dalam pengolahan data teks, pendekatan *Deep Learning* banyak diterapkan dalam bidang *Natural Language Processing* (NLP), yang bertujuan untuk memungkinkan komputer memahami dan menginterpretasikan bahasa manusia secara otomatis. Salah satu arsitektur *deep learning* yang berkembang pesat dalam NLP adalah Transformer yang diperkenalkan oleh Vaswani dkk. (2017). Transformer adalah mekanisme *self-attention* yang memungkinkan model untuk mempelajari hubungan antara setiap kata dalam sebuah kalimat tanpa bergantung pada urutannya. Transformer memiliki dua komponen, yaitu *decoder* digunakan untuk menghasilkan output teks seperti dalam terjemahan, dan *encoder* menghasilkan representasi kontekstual dari *input* teks (Darmawan, 2025).

Bidirectional Encoder Representations from Transformer (BERT) adalah salahsatu pengembangan *Transformer* yang menggunakan bagian *encoder* untuk

menangkap hubungan antar kata dalam teks secara efektif dan efisien. BERT diperkenalkan oleh *Google Research* melalui penelitian Devlin dkk., (2019). BERT merupakan model representasi bahasa yang dirancang sebagai *Transformer* dua arah (*bidirectional*), yaitu memproses setiap kata dengan mempertimbangkan seluruh konteks kalimat secara simultan dari kedua arah, baik dari kiri ke kanan maupun dari kanan ke kiri. Dengan demikian, representasi suatu kata dibentuk berdasarkan hubungan kata-kata sebelum dan sesudahnya, sehingga menghasilkan pemahaman makna yang utuh dan kontekstual. Pendekatan ini memungkinkan model menangkap makna yang bergantung pada konteks lebih akurat, karena interpretasi suatu kata tidak ditentukan secara parsial, melainkan berdasarkan keseluruhan struktur kalimat. Hal ini menjadi keunggulan model BERT dibanding model konvensional sebelumnya yang bersifat *unidirectional*, yaitu hanya memanfaatkan satu arah konteks, sehingga memiliki keterbatasan dalam menangkap makna yang kompleks.

BERT telah dikembangkan ke dalam berbagai model spesifik bahasa untuk meningkatkan pemahaman terhadap karakteristik linguistik masing-masing bahasa. Salah satu pengembangannya adalah IndoBERT (*Indonesian Bidirectional Encoder Representations from Transformer*), yaitu model BERT yang menjalani proses *pre-training* menggunakan korpus teks berbahasa Indonesia dalam skala besar, seperti *Wikipedia Corpus* dan *BookCorpus* dengan lebih dari 220 juta kata (Wilie dkk., 2020). Oleh karena itu, IndoBERT disebut *pre-trained language model*, yaitu model yang telah mempelajari pola umum bahasa Indonesia sebelum digunakan pada tugas spesifik. Pada tahap *pre-training*, IndoBERT dilatih secara *unsupervised learning* menggunakan *Masked Language Modeling* (MLM) dan *Next Sentence Prediction* (NSP). MLM melatih model untuk memprediksi token yang disembunyikan berdasarkan konteks di sekitarnya, sedangkan NSP melatih model untuk memahami hubungan antar dua kalimat. Melalui proses tersebut, IndoBERT mampu mempelajari struktur bahasa, hubungan antar kata, serta representasi semantik bahasa Indonesia, sehingga dapat digunakan kembali pada berbagai tugas NLP melalui proses *fine-tuning*.

Secara arsitektural, proses IndoBERT dimulai dari *preprocessing* untuk membersihkan bagian-bagian *tweet* yang tidak memiliki makna. Kemudian dilakukan tokenisasi, dimana isi *tweet* diubah menjadi unit-unit kecil berupa token menggunakan metode *WordPiece*. *WordPiece* memecah kata menjadi bagian-bagian kecil (*subword*) berdasarkan kosakata model untuk mengatasi permasalahan *out-of-vocabulary* (OOV). Melalui metode ini, kata yang tidak dikenali secara utuh tetap dapat direpresentasikan berdasarkan bagian penyusunnya (Song dkk., 2021). Selanjutnya, setiap token direpresentasikan dalam bentuk vektor melalui proses *embedding*. Representasi ini kemudian diproses pada dua belas lapisan *encoder Transformer*. Pada setiap lapisan *encoder*, dilakukan operasi *Multi-head Self-attention* untuk menangkap hubungan antar token dalam konteks kalimat, serta *Feed Forward Network* (FFN) untuk memproses informasi secara non-linier. Proses ini berulang pada setiap *layer*, sehingga menghasilkan pemahaman *tweet* yang semakin hierarkis. Untuk tugas klasifikasi, IndoBERT memanfaatkan representasi token khusus CLS yang berada di awal input. Representasi token khusus CLS ini menangkap informasi dari seluruh kalimat, yang kemudian digunakan sebagai *input* pada lapisan klasifikasi untuk menghasilkan prediksi (Boukabous & Azizi, 2022).

Setelah tahap *pretraining* selesai, model IndoBERT disesuaikan dengan tugas spesifik melalui proses *fine-tuning*. Pada tahap *fine-tuning*, ditambahkan lapisan klasifikasi multikelas (*classification layer*) pada bagian *output* model *pre-trained* IndoBERT untuk menghasilkan prediksi kelas emosi. Parameter hasil *pre-training* digunakan sebagai inisialisasi awal, kemudian seluruh parameter model pada lapisan *embedding*, *encoder Transformer*, dan lapisan klasifikasi tersebut kemudian diperbarui melalui proses pelatihan ulang (*fine-tuning*) menggunakan dataset *tweet* berbahasa Indonesia yang telah diberi label emosi. Proses *fine-tuning* memungkinkan model menyesuaikan representasi bahasa yang telah dipelajari sebelumnya dengan karakteristik data spesifik, sehingga meningkatkan akurasi dan kemampuan generalisasi model (Kannan & Kothamasu, 2020). Dalam penelitian ini, proses *fine-tuning* dilakukan pada data *tweet* berbahasa Indonesia terkait program MBG.

Penelitian oleh Delvin, dkk. (2025) menggunakan metode *Fine-tuning*

IndoBERT untuk menganalisis sentimen pada 10.000 data ulasan pengguna aplikasi Tiket.com di Google Play Store mampu mencapai akurasi sebesar 91%. Selanjutnya, penelitian oleh Haikal, dkk. (2025) menggunakan metode Fine-tuning IndoBERT untuk klasifikasi emosi multikelas (*anger, fear, happy, love, dan sadness*) pada 4.940 *tweet* yang berisikan teks informal bahasa Indonesia mampu memperoleh akurasi sebesar 94%. Hasil penelitian ini menunjukkan bahwa metode Fine-tuning IndoBERT efektif digunakan untuk klasifikasi emosi multikelas pada teks berbahasa Indonesia.

Penelitian ini bertujuan untuk menganalisis multikelas emosi publik terhadap program Makan Bergizi Gratis (MBG) menggunakan metode *fine-tuning* IndoBERT pada data *tweet* media sosial X. Klasifikasi emosi dilakukan ke dalam enam kategori berdasarkan Teori Paul Ekman, yaitu senang, sedih, marah, takut, jijik, dan terkejut, sehingga multikelas emosi tersebut lebih mampu menggambarkan dinamika emosi publik dibandingkan pendekatan emosi biner. Dengan memanfaatkan kemampuan IndoBERT dalam memahami konteks bahasa secara mendalam, penelitian ini diharapkan mampu menghasilkan model klasifikasi emosi yang akurat terhadap data *tweet* media sosial X. Hasil penelitian ini diharapkan dapat memperkaya kajian dalam bidang NLP, serta menjadi bahan evaluasi bagi pemerintah dalam memahami respon publik terhadap implementasi program MBG.

1.2. Rumusan Masalah

1. Bagaimana persepsi emosi publik terhadap Program MBG yang disampaikan melalui media sosial X?
2. Bagaimana performa model Fine-tuning IndoBERT dalam mengklasifikasikan emosi publik ke dalam enam kategori berdasarkan Teori Ekman?
3. Emosi apa yang paling dominan muncul pada respon publik terhadap Program MBG?

1.3. Batasan Masalah

1. Data yang digunakan berupa *tweet* berbahasa Indonesia pada media sosial

X yang berkaitan dengan Program Makan Bergizi Gratis (MBG) selama periode awal pelaksanaan, yaitu sejak 1 Januari 2025 hingga 31 Oktober 2025.

2. Seluruh data *tweet* yang memenuhi kriteria penelitian digunakan sebagai data penelitian dan dilabeli ke dalam enam label emosi berdasarkan teori Ekman, yaitu senang, sedih, marah, takut, jijik, dan terkejut.

1.4. Tujuan Penelitian

1. Menganalisis dan mengklasifikasikan emosi publik terhadap Program MBG pada media sosial X.
2. Menerapkan metode Fine-tuning pada model IndoBERT untuk klasifikasi emosi multikelas berdasarkan Teori Ekman.
3. Mengidentifikasi distribusi emosi paling dominan sebagai bentuk respon publik terhadap Program MBG.

1.5. Manfaat Penelitian

1. Manfaat akademik

Penelitian ini berkontribusi pada pengembangan ilmu khususnya dalam bidang *Natural Language Processing* (NLP) dalam menganalisis multikelas emosi publik.

2. Manfaat Praktis

Hasil analisis ini dapat dijadikan sebagai referensi oleh pemerintah untuk memahami persepsi publik lebih mendalam, sehingga membantu dalam menyusun strategi implementasi program yang efektif.